



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/842,358	04/24/2001	Pieter Vermeulen	SEN-112	6566
30869	7590	05/20/2004	EXAMINER	
LUMEN INTELLECTUAL PROPERTY SERVICES, INC.				
2345 YALE STREET, 2ND FLOOR				
PALO ALTO, CA 94306				
			ART UNIT	PAPER NUMBER
			2654	9

DATE MAILED: 05/20/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

# Office Action Summary

Application No.

09/842,358

Applicant(s)

VERMEULEN ET AL.

Examiner

Daniel A. Nolan

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☒ Responsive to communication(s) filed on 12 January 2004.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1-35 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-35 is/are rejected.
- 7) ☒ Claim(s) 6,7,10,16,21 and 35 is/are objected to.
- 8) ☒ Claim(s) 1-5,8,9,11-15,17-20 and 22-34 are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 24 April 2001 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on \_\_\_\_\_ is: a) ☐ approved b) ☐ disapproved by the Examiner.  
If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

## Priority under 35 U.S.C. §§ 119 and 120

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a) ☐ All b) ☐ Some \* c) ☐ None of:  
1. ☐ Certified copies of the priority documents have been received.  
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).  
\* See the attached detailed Office action for a list of the certified copies not received.
- 14) ☒ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).  
a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

## Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 8.
- 4) ☐ Interview Summary (PTO-413) Paper No(s). \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other:

### **DETAILED ACTION**

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

#### ***Response to Amendment***

2. The filing of 12 January 2004 was applied to the following effect:
  - The specification was changed as indicated and the objections are withdrawn.
  - The claims were changed as indicated and the objections are withdrawn as satisfied.The revised claims were examined on the merits.

#### ***Response to Arguments***

3. Applicant's claim for domestic priority under 35 U.S.C. 119(e) is acknowledged in view of the argument that the disclosure of the invention in the prior application and in the later-filed application is at least sufficient to comply with the requirements of the first paragraph of 35 U.S.C. 112.

***Claim Rejections - 35 USC § 102***

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. Claims 1, 4, 5, 9, 11, 15, 18, 19, 23-27 and 30-35 are rejected under 35 U.S.C. 102(b) as being anticipated by Matsumoto et al<sup>'362</sup> (U.S. Patent 5,673,362 A).

6. Regarding claims 1, 9 and 15, Matsumoto et al<sup>'362</sup> reads on the claims for *text-to-speech synthesis in a computer system comprising a server machine and a client machine* specific to claims 1 and 9 as follows:

- Matsumoto et al<sup>'362</sup> reads on the feature of *a computer system comprising a server machine and a client machine* (10 & 20 In figure 2)
- Matsumoto et al<sup>'362</sup> reads on the feature of *text-to-speech synthesis* (11→15 in figure 2)
- Matsumoto et al<sup>'362</sup> reads on the feature of *describing a finite number of possible acoustic units* (with the *preprocessing* of column 9 lines 33-37);

Art Unit: 2654

- Matsumoto et al<sup>362</sup> reads on the feature of *optimizing a compression method selected in dependence of the finite number of possible acoustic units* (column 11 lines 32-33);
- Matsumoto et al<sup>362</sup> reads on the feature of *compressing the finite number of possible acoustic units via the optimized compression method* (column 12 lines 40-56);
- Matsumoto et al<sup>362</sup> reads on the feature of *storing the finite number of possible acoustic units as compressed acoustic units in an acoustic unit database* (53 in figure 6) *accessible to the server machine* (20→50 in figure 6);
- Matsumoto et al<sup>362</sup> reads on the feature of *in the server machine, obtaining normalized text* (11 in figures 1-4 – see the *preprocessor* of column 9 lines 33-37) *and generating prosody data* (column 9 lines 43-44);
- Matsumoto et al<sup>362</sup> reads on the feature of *selecting from the acoustic unit database compressed acoustic units that correspond to the normalized text* (column 19 lines 23-38);
- Matsumoto et al<sup>362</sup> reads on the feature of *transmitting the prosody data and the selected compressed acoustic units from the server machine to the client machine* (20→40 in figure 6); *and*
- Matsumoto et al<sup>362</sup> reads on the feature of *decompressing transmitted acoustic units* (with decoding, 14 in figure 2) *and concatenating the selected decompressed acoustic units* (14→151 in figure 2) *in accordance with prosody data* (221 & 222 in figure 2 – see column 9 lines 43-44).

With respect to the similar features of claim 15, the invention of Matsumoto et al<sup>362</sup> for a speech synthesis system in which a plurality of clients and at least one voice synthesizing server are connected to a local area network reads on every feature of the claim in a client machine, a text-to-speech synthesis method comprising the following:

- Matsumoto et al<sup>362</sup> reads on the feature of receiving compressed acoustic units (26→13 in figure 2) corresponding to a normalized text from a server machine (11 in figures 1-4 – see the preprocessor of column 9 lines 33-37), the compressed acoustic units being selected from a predetermined number of possible acoustic units and compressed using a compression method selected in dependence on the predetermined number of possible acoustic units (25 in figure 2 – see column 10 lines 38-44);
- Matsumoto et al<sup>362</sup> reads on the feature of decompressing the compressed acoustic units to obtain decompressed acoustic units (with decoding 14 in figure 2); and
- Matsumoto et al<sup>362</sup> reads on the feature of concatenating the decompressed acoustic units (151 in figure 2).

7. Regarding claims 4, 18 and 29; the claims are set forth with the same limits as claims 1, 15 and 25, respectively. Matsumoto et al<sup>362</sup> reads on the feature of normalizing a standard text to obtain the normalized text (column 19 lines 23-38).

8. Regarding claims 5, 11 and 19; the claims are set forth with the same limits as claims 1, 9 and 19, respectively. Matsumoto et al<sup>362</sup> reads on the features of sending a

*standard text to the server machine (12→21 in figure 1) and, in the server machine, normalizing the standard text to obtain the normalized text (220-221 in figure 1).*

9. Regarding claims 23, 24 and 25; the claims are set forth with the same limits as claims 1, 9 and 19, respectively. Matsumoto et al<sup>362</sup> reads on the features of the *acoustic unit database the server machine in communication with the acoustic unit database; and the client machine in communication with the server machine (10 & 20 in figure 2), the acoustic unit database (224 in figure 1) for storing the predetermined number of possible acoustic units; means for enabling data transmission and communication among the acoustic unit database the server machine, and the client machine (224→26→13 in figure 2).*

10. Regarding claim 26, the claim is set forth with the same limits as claim 25. The features of the claim are the same as those found in claim 4 and the claim is rejected for the same reasons.

11. Regarding claim 27, the claim is set forth with the same limits as claim 25. The features of the claim are the same as those found in claim 15 and the claim is rejected for the same reasons.

12. Regarding claims 30, 31, 32, 33, 34 and 35; the claims are set forth with the same limits as claims 1, 9, 15, 19, 20 and 21, respectively. Matsumoto et al<sup>362</sup> teaches

Art Unit: 2654

the process of distributed or client/server voice synthesizing in such a way that would be inherently within the expected skill of an ordinary practitioner in the art of speech signal processing, to accomplish the following (with reference to column 1 lines 16-28):

- Matsumoto et al<sup>362</sup> teaches the inherent feature of providing a *computer-readable program storage device tangibly embodying a computer-executable program implementing the text-to-speech synthesis method* of claim 30 with this above reference.
- Matsumoto et al<sup>362</sup> teaches the inherent feature of a *computer-readable medium storing a computer-executable program implementing the text-to-speech synthesis method* of claim 31 with this above reference.
- Matsumoto et al<sup>362</sup> teaches the inherent feature of providing *computer-readable medium storing a computer-executable program implementing the text-to-speech synthesis method* of claims 32, 33, 34 and 35, all with this above reference.

### ***Claim Rejections - 35 USC § 103***

13. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to

Art Unit: 2654

consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

**Matsumoto et al<sup>'362</sup> & Akamine et al**

14. Claims 1, 9, 15 are rejected under 35 U.S.C. 103(a) as being unpatentable in view of Matsumoto et al<sup>'362</sup> (U.S. Patent 5,673,362 A) in view of Akamine et al (U.S. Patent 6,161,091 A).

15. Regarding claims 1, 9 and 15, Matsumoto et al<sup>'362</sup> reads on the claims for *text-to-speech synthesis in a computer system comprising a server machine and a client machine* as follows:

- Matsumoto et al<sup>'362</sup> reads on the feature of *a computer system comprising a server machine and a client machine* (10 & 20 In figure 2)
- Matsumoto et al<sup>'362</sup> reads on the limit of *text-to-speech synthesis* (11→15 figure 2)
- Matsumoto et al<sup>'362</sup> reads on the feature of *describing a finite number of possible acoustic units* (with the *preprocessing* of column 9 lines 33-37);
- Matsumoto et al<sup>'362</sup> reads on the feature of *optimizing a compression method selected in dependence of the finite number of possible acoustic units* (column 11 lines 32-33);
- Matsumoto et al<sup>'362</sup> reads on the feature of *compressing the finite number of possible acoustic units via the optimized compression method* (column 12 lines 40-56);

Art Unit: 2654

- Matsumoto et al<sup>362</sup> reads on the feature of *storing the finite number of possible acoustic units as compressed acoustic units in an acoustic unit database (53 in figure 6) accessible to the server machine (20→50 in figure 6);*
- Matsumoto et al<sup>362</sup> reads on the feature of *in the server machine, obtaining normalized text (11 in figures 1-4 – see the preprocessor of column 9 lines 33-37) and generating prosody data (column 9 lines 43-44);*
- Matsumoto et al<sup>362</sup> reads on the feature of *selecting from the acoustic unit database compressed acoustic units that correspond to the normalized text (column 19 lines 23-38);*
- Matsumoto et al<sup>362</sup> reads on the feature of *transmitting the prosody data and the selected compressed acoustic units from the server machine to the client machine (20→40 in figure 6); and*
- Matsumoto et al<sup>362</sup> teaches that the client machine synthesizes speech but does not describe the steps of that well-known process. Akamine et al, with the invention for a speech recognition-synthesis based encoding-decoding method, and speech encoding-decoding system, reads on the feature of *decompressing transmitted acoustic units (with decoding in column 3 line 60) and concatenating the selected decompressed acoustic units (with synthesizing column 3 line 63) in accordance with prosody data (column 3 line 64)*. It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Akamine et al to the device/method of Matsumoto et al<sup>362</sup> so as to transfer non-linguistic information such as a speaker's feeling.

**Matsumoto et al.<sup>362</sup> & Castello da Costa et al**

16. Claims 2, 12 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable in view of Matsumoto et al.<sup>362</sup> in view of Castello da Costa et al (U.S. Patent 6,230,130 B1).

17. Regarding claims 2, 12 and 22; the claims are set forth with the same limits as claims 1 and 9, respectively. Matsumoto et al.<sup>362</sup> does not speak to *streaming* as disclosed in the application for processing as received (page 9 lines 16-19).

Castello da Costa et al, with the invention of *scalable mixing for speech streaming*, teaches the feature where the *decompressing step and the concatenating step begin before all of the selected compressed acoustic units and the prosody data are received in the client machine* (column 1 lines 15-19).

It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Castello da Costa et al to the device/method of Matsumoto et al.<sup>362</sup> to allow the benefit of concurrent processing so that the time the processing is completed is shortly after the last segment is received.

**Matsumoto et al'<sup>362</sup>, Akamine et al & Van Kommer**

18. Claims 3, 14, 17 and 28 are rejected under 35 U.S.C. 103(a) as being unpatentable in view of Matsumoto et al'<sup>362</sup> in view of Van Kommer (U.S. Patent 6,678,659 B1).

19. Regarding claims 3, 14, 17 and 28; the claims are set forth with the same limits as claims 1, 9, 15 and 25, respectively. Matsumoto et al'<sup>362</sup> does not mention *caching*. The invention of Van Kommer for *voice information dissemination over a network using semantic representation* reads on the features of *caching a number of frequently used uncompressed acoustic units in a cache memory of the client machine* (column 3 lines 37-41) and *concatenating the decompressed acoustic units with at least one of the uncompressed acoustic units* (column 4 line 65 – column 5 line 10). It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Van Kommer to the device/method of Matsumoto et al'<sup>362</sup> so as to limit the number of necessary conversions.

**Matsumoto et al'<sup>362</sup> & Malsheen et al**

20. Claims 8, are rejected under 35 U.S.C. 103(a) as being unpatentable in view of Matsumoto et al'<sup>362</sup> in view of Malsheen et al (U.S. Patent 4,979,216 A).

Art Unit: 2654

21. Regarding claims 8, the claim is set forth with the same limits as claim 1.

Matsumoto et al<sup>362</sup> does not speak to *dividing acoustic units* but Akamine et al teaches the step of *dividing each of the possible acoustic units into sequences of chunks of equal duration* (column 7 lines 55-58 & 60-64).

Neither Matsumoto et al<sup>362</sup> nor Akamine et al mention *frequency parameters*. Malsheen et al, with the invention of *text to speech synthesis system and method using context dependent vowel allophones*, reads on the step of *describing frequency composition of each chunk with a set of parameters* (column 2 lines 34-55, listing the requirements, strengths and weaknesses, so not teaching against the technique). It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Malsheen et al to the device/method of Neither Matsumoto et al<sup>362</sup> & Akamine et al so as to provide more human-sounding speech.

**Matsumoto et al<sup>362</sup> & Levine et al**

22. Claims 13 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable in view of Matsumoto et al<sup>362</sup> in view of Levine et al (U.S. Patent 5,886,276 A).

23. Regarding claims 13 and 20, the claims are set forth with the same limits as claim 15. Matsumoto et al<sup>362</sup> does not speak to *quality* as a factor in selecting the coder. The invention of Levine et al for *multi-resolution scalable audio signal encoding* reads on the feature of *selecting parameters of the compression method to minimize the*

Art Unit: 2654

*amount of data transmitted to the client machine while maintaining a minimum acoustic quality for each of the possible acoustic unit* (column 12 lines 1-3) ... It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Levine et al to the device/method of Matsumoto et al<sup>362</sup> so as to match the coding method with the transmission requirements.

### ***Allowable Subject Matter***

24. Claims 6, 7, 10, 16, 21 and 35 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

25. The following is a statement of reasons for the indication of allowable subject matter:

- The present invention is directed to *distributed speech synthesizing*.
- Regarding claim 6, the closest prior art of Matsumoto et al<sup>362</sup> has the client synthesizing speech signals furnished by the server, but without further parameters, so the feature of selecting parameters of the compression method utilizing a directed optimized search to minimize the amount of data transmitted between the server machine and the client machine was neither anticipated nor was it found in obvious combination in the prior art of record.

Art Unit: 2654

- Regarding claims 10, the closest prior art of Matsumoto et al<sup>362</sup> discloses the server *generating prosody data corresponding to the normalized text and transmitting the prosody data to the client machine* but fails to anticipate or render the above underlined limitations obvious.
- Regarding claim 16, related to the findings for claim 10, above, the feature that the synthesizing or *concatenating decompressed acoustic units depending on the prosody data* from the server was not found to be anticipated nor was it found in obvious combination in the prior art of reference.
- Regarding claim 21, the closest prior art of Akamine et al associates quality with searching but without anticipating or rendering obvious alone or in combination the feature of *utilizing an optimized search directed by an acoustic metric that measures minimum acoustic quality*.
- Claims 7 and 35 would be allowed because they depend on claims that were found to be allowable.

### ***Conclusion***

26. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

27. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- Ostrowski<sup>'730</sup> (U.S. Patent 4,130,730 A) voice synthesizer.
- Stentiford et al (U.S. Patent 5,384,701 A) language translation system.
- Castello da Costa et al (U.S. Patent 6,230,130 B1) scalable mixing for speech streaming.
- Tei (U.S. Patent 5,943,648 A) speech signal distribution system providing supplemental parameter associated data.
- Matsumoto et al<sup>'628</sup> (European Patent 542628 A2) client/server speech synthesis system.
- Ostrowski<sup>'150</sup> (U.S. Patent 4,470,150 A) voice synthesizer with automatic pitch and speech rate modulation.

Art Unit: 2654

28. Any inquiry concerning this communication or earlier communications from the Examiner should be directed to Daniel A. Nolan at telephone (703) 305-1368 whose normal business hours are Mon, Tue, Thu & Fri, from 7 AM to 5 PM.

If attempts to contact the examiner by telephone are unsuccessful, supervisor Richemond Dorvil can be reached at (703)305-9645.

The fax phone number for Technology Center 2600 is (703)872-9314. Label informal and draft communications as "DRAFT" or "PROPOSED", & designate formal communications as "EXPEDITED PROCEDURE". Formal response to this action may be faxed according to the above instructions,

or mailed to: Mail Stop AF (or CPA, etc. – see Official Gazette, 04 November 2003)  
P.O. Box 1450  
Alexandria, VA 22313-1450

or hand-deliver to: Crystal Park 2,  
2121 Crystal Drive, Arlington, VA,  
Sixth Floor (Receptionist).

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to Technology Center 2600 Customer Service Office at telephone number (703) 306-0377.

Daniel A. Nolan  
Examiner  
Art Unit 2654

DAN/d  
May 10, 2004

  
RICHEMOND DORVIL  
SUPERVISORY PATENT EXAMINER